



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2008

---

## **Role of normativity for the explanation of Norm-Conformity**

Herfeld, Catherine

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-142307>

Conference or Workshop Item

Published Version

Originally published at:

Herfeld, Catherine (2008). Role of normativity for the explanation of Norm-Conformity. In: Graduate Philosophy Conference on Normativity, Amsterdam, 29 August 2008 - 30 August 2008. s.n., 71-104.

Proceedings of the  
Graduate Philosophy Conference on Normativity  
Amsterdam, August 29 & 30, 2008

# Contents

Naturalism and Normativity: Sellars' 'Janus-Faced' Space of Reasons <i>Dr. Jim O'Shea (University College Dublin)</i>	<b>1</b>
Assessment-sensitivity and the Naturalistic Fallacy <i>Joshua Beattie (University of California, Berkeley)</i>	<b>15</b>
The Sources of Epistemic Normativity <i>Anthony Booth (Utrecht University)</i>	<b>55</b>
Role of Normativity for the Explanation of Norm-Conformity <i>Catherine Herfeld (Witten/Herdecke University)</i>	<b>71</b>
Naturalism and Normativity <i>Christos Kyriacou (University of Edinburgh)</i>	<b>105</b>
The Role of Explicitness in Assessing Practical Reasoning <i>Andrea Lechler (University of Reading)</i>	<b>127</b>
Perception, Causation and Justification <i>Graham Peebles (University of Glasgow)</i>	<b>149</b>
Moral and Rational "Oughts": The Distinction between 'Demanding' and 'Recommending' Normativity <i>Peter Schulte (University of Bielefeld)</i>	<b>161</b>
Only a Chapter of Psychology? The Defense of Normative Naturalism in Epistemology <i>Giedre Vasiliauskaite (Erasmus University Rotterdam)</i>	<b>169</b>

# Role of Normativity for the Explanation of Norm-Conformity

*Catherine Herfeld (Witten/Herdecke University)*

catherine.herfeld@uni-wh.de

## I Introduction

Nowadays it is widely accepted that social norms play a significant role in the explanation of cooperative behaviour and the maintenance of social order. People use social norms such as fairness, reciprocity, and promise-keeping to interact with each other (Henrich, 2000). While explaining human behaviour, economists have realized that the concept of social norms can contribute to the understanding of (economic) behaviour (Basu and Weibull, 2003; Conlin et al., 2003; Fehr et al., 1998; Kahneman et al., 1986; Lindbeck et al., 1999). Despite the problem that social norms are a) hard to measure and b) only vaguely understood in terms of their actual influence on people's behaviour, economists increasingly integrate social norms into their frameworks (Akerlof, 1982; Bicchieri, 2006; Conlin et al., 2003; Fehr et al., 1998). The framework I am mainly referring to here is the theory of rational choice (RCT).

When explaining action by using RCT, economists focus on observable behaviour. Such observed behaviours are then attributed to desires and beliefs as their causes, usually at the expense of other possible motivations. Models are based on the behavioural assumption of instrumental rationality and agency is determined by self-interest. If the agent is rational, his sole motive for action is assumed

to be personal welfare-improvement (Davis, 2004; Sen, 1985).<sup>1</sup> Consequently, the performed action is seen as a form of optimizing behaviour, taken to maximize individual utility and to satisfy personal preferences (Elster, 1989a). Transferring this conception unto the issue of social norms draws a picture of an internal motivating source for norm-conformity that is purely instrumental in character (Opp, 1999; Coleman, 1990).

Empirical results of ultimatum-, dictator-, or public-good-games show that economists cannot (always) predict cooperative behaviour. Instead, individuals will often follow social norms of cooperation at the expense of self-interest (Bicchieri, 2006; Fehr and Fischbacher, 2004). Thus, for the economist, sometimes the motivational force remains obscure. Now there are at least two ways of accounting for these ‘behavioral anomalies’: One way is to simply reduce them to irrational behaviour. This does not seem to be reasonable (Dawes and Thaler, 1988). If all actions are reduced to mere self-interest, even defined in a very broad sense (Opp, 1999), RCT becomes analytical. In this case, its status as a theory is generally questionable. A second way would be to refer to preferences for social norms as a solution for this explanatory dilemma. But, as claimed in this paper, to make social norms work, some additional motives are required to be in place. Different motives such as commitment play a crucial role in motivating norm-conformity itself and hence in explaining cooperation. This is because norm-conformity is often motivated by commitment. In particular, commitment motivates conformity to social norms of cooperation, *i.e.* to the norm of fairness, reciprocity and promise-keeping, whereby cooperative norms are used to interact with others. They enable or enhance cooperation among members of a given group or society (Henrich, 2000). Social norms are key elements in the provision of public goods. They prescribe solutions to social dilemma situations, *e.g.* to collective action problems that would otherwise exacerbate the provision and the maintenance of social order. If conformity to these norms is motivated by commitment, commitment is then a crucial motive in enabling cooperation.

---

<sup>1</sup>Welfare here is the satisfaction of preferences and thus understood in the neoclassical sense. This view is also called “preference satisfaction view” (Hausman and McPherson, 2006, p. 120). It implies that maximizing utility leads to personal welfare improvement, although both notions are not considered as the same and not synonymous from the philosopher’s point of view.

To overcome the explanatory shortcomings of the behavioural foundations of economic theory by preventing RCT from analyticity and in order to find adequate explanations for norm-conformity, we have to distinguish between different types of motivation which must be investigated beyond self-interest. To differentiate these different types of motivation, it is useful to analyse the normativity involved in norm-conformity. This will create an understanding of the role normativity plays in the explanation of norm-conformity. To tackle the issue, this paper is concerned with finding an answer to the question ‘what kind of normativity is involved in norm-conformity, *i.e.* why do people think that they ought to follow a social norm?’

## II Social Norms, Commitment, and Rational Choice Theory

Norm-conformity is often explained by economic frameworks, such as RCT. However, such frameworks are mainly based on the assumption of self-interested agents. If a person is modelled to act in his or her self-interest, this entails that he or she is driven by a reason consisting of a desire to achieve a certain end and a belief, that a specific action will achieve this end. Being motivated in this way implies that the performed action is an instrument for reaching a further end and therefore excludes the view that an action can be performed for its own sake. Applying this view to norm-conformity, the assumption that an agent is motivated by self-interest implies that the agent’s conformity serves solely as an instrument for him to achieve a certain further end. Holding a reason of obedience to a social norm because one aims at some further end, whereas the conformity would serve as a means to achieve that end, is what I will call holding an ‘instrumental reason’ for norm-conformity. The normativity involved in such norm-conformity, *i.e.* the reason why people think they ought to follow a social norm, is purely instrumental in character. In contrast, the motive of commitment provides people with reasons for obedience that are fundamentally different in nature. As I will argue in this paper, the motive of commitment provides practical reasons that are “independent of the gains and losses for the person in case he or she acts on that reason” (Pauer-Studer, 2007, p. 75). This is because these commitments are not based on the agent’s desires (Schmid, 2007). They do not express what the agent *wants*. Rather, if a commitment presents a reason for action this is because commitments, as Searle puts it, create

*desire-independent reasons* for action (Searle, 2001). If we accept commitment to be a relevant motivation for people to follow social norms, the normativity involved in the resulting norm-conformity differs from that of the instrumentally-driven norm-conformity described above. This difference in the type of normativity involved in conformity then depends on the motivation of the people.<sup>2</sup> As I argue in this paper, this is why commitment as a motivational force cannot be analysed by a framework based on the idea of instrumental action and the motive of self-interest. If this is the case, classical economic theory cannot explain all types of norm-guided behaviour.

There are three ways to deal with the ensuing explanatory dilemma: Firstly, many economists simply dismiss the motive of commitment as an important motivating force for norm-conformity. But this is problematic, because commitment seems to offer a good explanation of different normative phenomena such as work motivation, solidarity bonds *etc.* (Pauer-Studer, 2007). And it provides a useful account of cooperation observed in experimental settings (Kerr et al., 1997).<sup>3</sup> A second possibility is to ignore the special nature of commitment and just assume that people have a preference for commitment; these approaches explain anomalies with as-if models based on long-term calculations and self-interest. This seems plausible at first sight. Economists have highly elaborated upon their concepts of self-interest and preferences thanks to the openness / flexibility of the utility-function. However, because all the motives covered by economic models are ultimately traced back to self-centred agency, which fundamentally contradicts with what is understood as an action motivated by commitment, this does not present a solution either. Additionally, defenders of this view are confronted with the objection of economic explanations then becoming merely analytic. Assuming a preference for commitment to explain committed behaviour seems to lead to a tautological explanation which provides very little information. The *explanandum*, *i.e.* the event or fact to be explained is at the same time used as part of the *explanans*, *i.e.* the

---

<sup>2</sup>Whether or not this leads to relativism of normativity will not be discussed here in greater detail but will be, at least indirectly, referred to later in the paper.

<sup>3</sup>I will not tackle the issue of providing empirical evidence for the existence of a motive of commitment in this paper. However, I am aware that this is a very delicate and important issue where a lot more work needs to be done. Interesting empirical literature is provided by Kerr et al. (1997).

mechanism which does the explanation. A third possibility open to economists is to take commitment seriously as a motivational force, to admit the explanatory shortcomings of the current framework and try to adapt it. In this case, the assumption of individual-welfare maximization and self-interested agency would have to be relaxed or even rejected (Sen, 1982). I am going to argue for this last possibility.

### III Four questions

To develop my argument I shall address the following four questions:

- I. How does classical economic theory explain why people think they *ought* to follow a social norm, *i.e.* what makes individual norm-conformity normative from the economist's perspective?
- II. What is the nature of this 'ought' or the type of normativity involved in norm-conformity and how does this 'ought' or type of normativity relate to motivation?
- III. What is this 'ought' or type of normativity grounded upon? Does it differ between people/cultures and, if so, why?
- IV. What role do motives and normativity play in explaining norm-conformity?

By addressing each of these questions in turn, I hope to show the methodological implications of analysing the normativity involved in norm-conformity. I shall leave the conceptual implications for later analyses.

#### I.

To address the first question of how classic economic theory explains why people think they *ought* to follow a social norm, one must first understand how actions in general and norm-conformity in particular are explained by economists and how they define social norms. Looking at the key features of social norms provides us with an understanding of the motivation people are assumed to have when conforming to social norms.

Why people act in the way they do is a difficult question. We cannot observe people's motivation for performing a certain action.



What goes on in people's heads is often taken to be inaccessible. People's minds are black boxes for the economist. However, a lot of empirical research has been done to find out why people conform to social norms. In general, the economist's explanation is based upon what he generally assumes to be characteristic for a motive of action. This view of motivation is based on the concept of intentional action used in folk psychology. An intentional action is a piece of behaviour which derives from an intention of the individual exhibiting it (Elster, 1994). This is what distinguishes action from mere behaviour, namely that it is done with a purpose or reason (*e.g.* Anscombe, 1963; Rosenberg, 1995). This means that every individual, when performing an action has a reason to do so.

Reasons are modelled as being provided by mental states, *i.e.* the desire to do something and the belief that a certain action serves as a means to satisfy this desire (Elster, 1994; Bicchieri, 2006; Hausmann, 2005). The motive for action is thus taken to be identified by providing a reason for an action. In this line of thought, the reason for an action represents the intention of the agent. If the person acts rationally, his intention finally motivates, *i.e.* causes, the action. Thus, by giving the reason for an action, the motivational structure of an action is seen as being causal in nature (Davidson, 1963). And the action of the agent is always assumed to be the intended action, based on what the agent desired and what he believed to be the means for fulfilling the desire.<sup>4</sup> This implies that the reason given by an intentional explanation is assumed to be the reason which ultimately motivated the agent to perform the action.<sup>5</sup> In this sense, this view supposes that an action is always done for a purpose; it is purely instrumental in nature.

In the economic language, the desires of the person are translated into a preference-set. The beliefs a person holds are taken to be the full set of information available to the agent. Given the end the agent aims to achieve, his preferences and information motivate him to perform an action. The action in question is seen as the

---

<sup>4</sup>I will not consider exceptions such as *akrasia* here.

<sup>5</sup>Note that it remains opaque and also uninteresting what the ultimate reason of a person is. The person might not be able to identify it and thinks he acted because of reason X but what really but unconsciously motivated the agent was reason Y. Or there might be a chain of reasons which jointly caused the action in question (Rosenberg, 1995).

best means to achieve his end (which is mostly the satisfaction of his personal preferences). Assuming this motivational structure, the action performed can be explained causally.

Preferences and beliefs are required to be internally and mutually consistent to call a choice instrumentally rational (Elster, 1994). However, consistency alone does not imply any kind of normativity of the agent's intentions in the sense that desires, the beliefs or the end pursued have valuable content for the agent or that the end is perceived as *right* in a more profound sense (Raz, 2005) — *right* in the sense that the agent believes or values it as a good or right thing, independent of what he is inclined to do or what would satisfy his preferences. There is also no space for the agent considering the action itself as worth pursuing, independently of his personal preferences.<sup>6</sup> What the agent evaluates or chooses as the right action goes back to the agent's ordering of his preference set. The agent orders his preferences on a scale. The measure is 'utility' and the ordering is represented by the utility offered by different (material or immaterial) goods. According to this order, the agent evaluates the outcomes of his actions; he chooses an action and acts accordingly. The judgement is assumed to be exclusively based on the criterion or principle of utility maximization. Because the outcome of the actions are ideally the fulfilment (or exceeding) of the actor's desires, the satisfaction of desires is assumed to always be the person's main end and the action is therefore said to be in his or her self-interest (Heap Hargreaves et al., 1992). In short, "desires [or preferences] are the unmoved movers" for human action in economic theory (Elster, 1989b, p. 4). The limited empirical basis for unobservable preferences or desires does not present a problem in this context. Instead, it is simply claimed that the observed action reveals the preference. And preferences reveal the motivations involved. In this understanding, different preferences (re-)present different motives; and preferences are given to economic theory.<sup>7</sup>

---

<sup>6</sup>To say that somebody prefers one state of affairs X to another Y "is to describe a mental state of that person which disposes her to choose actions which lead to X rather than actions which lead to Y" (Sugden, 2000).

<sup>7</sup>This relates to the idea of revealed preferences where it is said that the action taken reveals the preference which motivated the action (Samuelson, 1938; Sen, 1977). This is not necessarily true. Somebody who is going to church might be acting on purely selfish terms but reveals the same preference as a firm believer. However, there is still an explanation missing why some people are motivated

If one transfers this general picture of action-explanation outlined above to the case of social norms, conformity is seen as purely instrumental. An instrumentally rational agent has a prudential or — as I would call it — ‘instrumental reason’ to obey the norm because norm-conformity presents the best strategy in this particular situation, given the strategies of all other players involved.<sup>8</sup> Thus, the economist has a clear response to question I: norm-conformity is normative for an instrumentally rational agent because it is — in one way or another — in his self-interest to do so. In the following I will talk about ‘instrumental normativity’ in this context. Does this view of ‘instrumental normativity’ really fit the concept of what we think a social norm is or represents? Let us first clarify what is generally understood among economists to be a social norm.

Within various ontological views and explanatory accounts, there is no unified definition of a social norm used by social scientists (Hechter and Opp, 2001; Stout, 2001; Okruch, 1999). However, there are some key features or properties generally considered to be common to social norms. First, social norms are characterized as informal rules of behaviour “that people follow for some reason other than the fear of legal sanctions” (Stout, 2001, p. 6). This means that they are not enforced by a formal agency in contrast to formal institutions, *e.g.* legal norms. Instead, they are enforced by informal sanctions executed by third parties, for example through social forces such as the exclusion from a social group by its members (Mantzavinos, 2001).

The informal character of social norms leads to the second feature, their social aspect. They are social in the sense that they are assumed to exist in every society, every subgroup of society and any kind of organisational structure (Kirchgässner, 2006). In economic approaches, social norms are often assumed to be localized in people’s preference set. The agent is modelled to have a preference for a social norm. This can be either by being (a) accepted by the members of a group or society or (b) through the internalization of the social norm by the individuals through the process of socialization

---

differently than others (Kurzban and Houser, 2005).

<sup>8</sup>Bicchieri defines a prudential reason for action as “if you have a goal X and the best available means to attain X is a course of action y, then you ought to adopt Y” (2006: 14). This is similar of having what I called ‘instrumental reason’ as an instrumentally rational agent.

(Hechter and Opp, 2001; Fehr and Fischbacher, 2004). In (a), the social norm is generally realized and consciously followed by the people, *i.e.* conformity involves choice on the part of the agent, based upon a preference for the end for which norm-conformity is the best means. Thus, social norms serve to as means for the individual to achieve or satisfy certain ends that *directly* affect its personal welfare. In case (b), the social norm can be followed unconsciously. Here, the social norm itself is considered to be the dominant preference of the agent. Thus, internalized norms *indirectly* affect the personal welfare of the individual through (positive or negative) emotions or external but informal (positive or negative) sanctioning.

There has been plenty of discussion about this ‘narrow’ behavioural basis of economic theory when using the ‘preference-approach’ to explain norm-conformity. The discussion is mainly centred on the assumption of the self-centred or selfish agent. It is claimed that many possible motives cannot be accounted for in this framework (Sen, 1985). As a possible solution economists invent a very broad notion of self-interest. The concept of preference has been extended to all kinds of tastes that people can have: material, immaterial, pro-social/altruistic/conditional, emotive, *etc.*<sup>9</sup> Thereby, self-interest becomes definitional to actions and social norms are also allowed to be motivated by altruism or other-regarding preferences. Even in these cases, conformity can be explained by classical economic theory as long as individual choices are made consistently with people’s set of preferences (*e.g.* Bicchieri, 2006; Stout, 2001; Coleman, 1987; Elster, 1989a; Woodward, 2009). But what matters for the argument in this paper is that — in these explanations outlined above — all action based on preferences always goes back to personal welfare-improvement and self-interest. Why is that problematic?

---

<sup>9</sup>The economic framework is used to accommodate 1) all types of social concerns, altruism, emotions, and informal sanctioning (reputation, esteem, social disapproval, disesteem, *etc.*). It 2) accommodates completely unaware, automatic and blindly norm-conformity. It 3) can explain social norms as means to internalize externalities such as for example littering (*e.g.* Coleman, 1987). 4) Norm-conformity is considered as instrumentally rational when motivated by direct sanctions and indirectly by emotions. And rational choice theory accounts 5) for norm-guided behaviour based on the idea of conditional cooperation, thus when we believe that other people conform to social norms and also expect us to do so in a specific situation.

Let's take the case of an altruistic action. To explain altruism, the economist equips the agent with an altruistic preference. Altruism can be defined as "costly acts that confer economic benefits on other individuals" (Fehr and Fischbacher, 2003, p. 785). The agent acts altruistically because enhancing other people's welfare is what he is aiming at, *i.e.* his preference. His ultimate desires, *i.e.* preferences, are completely other-regarding (Bicchieri, 2006). But is this picture of the economist really what we understand to be altruistic? The agent prefers acting altruistically and caring about other people's welfare *because* his welfare depends on the outcome. His welfare is ultimately affected through present or future immaterial rewards. By acting altruistically, the agent improves his personal welfare to the extent of, for example, 'feeling good' or being socially esteemed. Altruism lies in his self-interest because of other than altruistic reasons, *i.e.* it is not his intention to be an altruist in the first place. It is simply the outcome of an action from which he benefits. This means that he does not care about the intrinsic value of altruism in the first place. He does not judge altruism itself as the right thing to do, as there is no judgement involved independent of personal welfare-improvement. It is only *because* he has a preference for altruism, which allows him to ultimately improve his personal welfare that he acts in the way that he does. Preferring altruistic behaviour is contingent and conditional on the outcome received by acting altruistically. But wouldn't we expect a Mother Theresa to be somehow unconditionally altruistic? Do we not think that, because she judges the intrinsic property of an altruistic action as being the right thing to do, that, if this judgement is based on her values, the performance of altruistic action becomes unconditionally necessary for her, *i.e.* independent of any rewards? And isn't this because she committed herself to certain religious values of a community which can be represented and implemented by the performance of an altruistic action? One could argue that an act may be utility-maximizing for a person because it satisfies the person's preferences yet at the same time reduces the personal welfare. Especially in the aforementioned example it seems hard to see how an unconditional and altruistic act benefits Mother Theresa, *i.e.* her level of personal welfare, in any real sense, although we could at the same time say that it satisfies her altruistic preference. Could we not simply say that the action in this case is consistent with utility-

maximization yet welfare-decreasing acts? This argument is valid when one considers welfare and utility-maximization to be fundamentally different. And we already made clear that we do not, as formal RCT equals them (Hausman and McPherson, 2006). The point I want to make here is that the concept of preferences is often misleading when it comes to explaining norm-conformity. Although it does not completely dismiss the normative dimension of agency, it gives a particular and overly narrow interpretation of this dimension. The capacity of a person to accept obligations and duties as motivating, to undertake judgements about right or wrong, good or bad, independent from immediate and self-centred desires, the capacity to base or adjust one's actions on these judgements and the capacity of self-determination/self-formation and self-control cannot be accounted for when taking the instrumental view of normativity that RCT implies.

As we have seen, when reducing norm-conformity to actions that are purely instrumental in character, the normativity involved in such actions is itself instrumental when the improvement of personal welfare is always the ultimate aim. This view contradicts the normative dimension of agency.<sup>10</sup> But this normative dimension seems to matter in the context of norm-conformity. That norms are often seen as prescribing how we ought to behave in certain situations seems to be left out by the economists view. This leads to the question II, which I want to address, namely what is the nature of this 'ought' or the type of normativity involved in norm-conformity?

## II.

The underlying problem of explaining norm-conformity with tools from economics is that norm-conformity seems to involve an instrumental *and* a normative aspect. A quick look over the literature allows us to pick out overlapping notions such as 'expectations', 'obligatory', 'value-judgements', 'right', or 'wrong' and 'authoritative' when it comes to defining social norms (Anderson, 2000; Bicchieri, 2006; Coleman, 1987; Elster, 1989a; Hechter and Opp, 2001). These normative notions suggest a type of normativity attached to

---

<sup>10</sup>An exception is the position of ethical egoism, *i.e.* the normative ethical position that moral agents ought to only do what is in their own self-interest, can be considered as normative dimension of agency.

norm-conformity that is different from the consistency-requirement postulated by RC-models. Social norms are not regarded as mere regularities in people's behaviour which can be observed.<sup>11</sup> Neither are they considered to be completely contingent in the sense that people's conformity varies fundamentally depending on their expected utilities in each situation.<sup>12</sup> Instead, social norms can be seen as more or less stable "*prescriptions* that establish how one *ought* to behave" within a community (Woodward, 2009, p. 33, my italics). They are defined as "standards of behaviour that are based on widely shared beliefs how individual group members *ought* to behave in a given situation" within the community (Fehr and Fischbacher, 2004, p. 185, my italics). They are said to "reinforce certain patterns of behaviour ... by representing these patterns as desirable or *obligatory*" within a community (Pettit, 1990, p. 725, my italics). In this sense, a social norm is a rule prescribing or proscribing behaviour. — It presents a statement of the form of a conditional, hypothetical imperative "in situation X, you *ought* to do Y" (they are conditional only on the circumstances and not, as instrumental rationality assumes, exclusively on outcomes) or of an unconditional, categorical imperative 'you should do X' (Elster, 1989a, p. 98, my italics).

People accept the social norm as a standard of evaluation for the behaviour on a collective level. Thus, these standards of right and wrong, *i.e.* the social norms themselves, have to be at least 'passively' approved of, if not actively ratified, enforced and followed as well as implemented by the individual.<sup>13</sup> The important point here is that such a collective acceptance requires a shared understanding of what is regarded as socially appropriate/inappropriate among the

---

<sup>11</sup>Bicchieri contrast social norms with descriptive norms such as fashions and fads, whereas the latter are mere regularities which can be observed. According to Bicchieri and in line with my argument, what classifies collective behaviour as a descriptive or a social norm is a) the motives of the people involved and b) the expectations of the people within the community (Bicchieri, 2006).

<sup>12</sup>There is a lot of empirical evidence that people's conformity varies from situation to situation but generally, social norms are often observed to be stable over time.

<sup>13</sup>'Passively approved of' in this context means, that the existence of a norm is accepted even if people avoid situations where the conformity to a social norm would have been expected, like the example of the Ik people shows (Turnbull, 1972). Thus, even if people do not follow the social norm, its existence is common knowledge and thus not denied and can therefore be assumed to be passively accepted.

individuals within a community (Krupka and Weber, 2008). But is it not that such a shared understanding of the required behaviour needs to have a basis for agreement, something like a standard, a principle or rationale on which an evaluation can be based? Such a basis is normally provided by certain fundamental evaluative principles or values of which the individual approves. Because the individual approves of these values and principles, he uses them to evaluate and judge actions as the right/wrong thing to do and finally perform the action. I argue that these standards and principles shape the individual's motives; they become part of what motivates the individual in performing a certain action. And this is why motives matter when it comes to explanations of norm-conformity. But how exactly do they shape the agents motives or even provide him with a motive for action?

We have seen so far that social norms can be either defined as means to personal welfare-improvement; or they can be defined as standards for the evaluation of behaviour within a community. Both views require an evaluation of alternative courses of actions that is based on an evaluative standard for deciding which course to take. But in the two cases the individual is assumed to base his choices on different evaluative principles. In the first case, the evaluation is assumed to be based on the utility-principle. The individual only takes into account the satisfaction of his personal preferences as the main aim. In the second case, the individual bases his evaluation on a standard which is external to the agent, which is set or prescribed by the community. In this case, the individual *and* the social perspective matter for the agent's choice of a course of action. Besides taking his personal preferences into account, the individual has to approve or at least accept this evaluative standard or principle of the community. It requires a shared understanding of what is considered as right and wrong among the individuals within a community. One could argue that such a shared understanding is not necessary and that norm-enforcement is generally secured by informal sanctions (Bicchieri, 2006; Coleman, 1987; Mantzavinos, 2001). And it is true that, as long as we do not live in a dictatorship, punishment admittedly exists to align people's behaviour to what the community considers as right or wrong. But it cannot uphold permanent and constant norm-conformity without any approval at the individual level, or at least from a group of reasonable size within the society.



Rather, people approve/disapprove of something which they value and judge as right or good, something which is in line with what they are convinced of. Therefore, shared values matter to enforce social norms.

Let's come back to the question of how people's principles and standards of evaluation relate to their motives for norm-conformity. The gap between standard or evaluation for judging an action as right or wrong, the agents' motivation to perform an action he considers as the right thing to do, and the actual performance of an action can be closed by introducing the idea of commitment. Before doing so, it is important to note that especially social norms of co-operation are operative by supporting value-laden patterns such as promise-keeping and truth-telling over a period of time. Individuals and groups transform their principles and value-orientation into their individual and collective actions. Social norms such as fairness, reciprocity, and promise-keeping align people's behaviours with their values and principles. Norm-conformity is a means for an agent to implement the values and principles of that individual and/or at the same time of the group or society that he is part of. Social norms can therefore fill a gap between values and principles on the one hand and conduct on the other (Wallace et al., 2004). Values and principles are at the same time constitutive for the self-conception of groups, individuals, and their self-understanding as autonomous agents. Thus, they play a role in supporting certain virtues (Elster, 1989a).

This is important because it shows the link between people's judgements and their motivation. In the case of social norms, people's conformity to them can be one way of revealing what the individual considers as being the right thing to do in a situation. In this case, the individual approves of what is seen to be right in the society. But, beyond mere approval, the individual can identify with certain values and principles that are embodied in social norms. The identification with values and principles generates commitment to them for groups and their members. In this case, one could say that people feel committed to these values and principles. And people think they ought to conform to a social norm because these represent their values and principles with which they identify apart from the fact that they are also considered to be right by the rest of the society or the members of the group. And because the individual has committed

to these values and principles, he himself judges the conformity to a social norm as an action which is the right thing to do. Conforming to them makes it possible to reveal the agent's commitments. It represents his values and principles and thereby allows for living up to his standards. Thus, the social norm itself becomes a standard of evaluation that is based upon and is consistent with what we generally judge as being good or bad, right or wrong. Thus, people's norm-conformity is motivated by commitment to these values and principles in the first place. Commitment provides the agent with a reason for norm-obedience; this is how his commitment motivates him to conform to a norm. But this reason is not instrumental in nature, because the agent's judgement in these cases is not based on the principle of instrumental rationality. The agent's performance of the action, *i.e.* his conformity to the social norm is independent of his personal situation; it is independent from his self-interest and his preferences. The appropriate motivation which captures this idea is his commitment as one possible motive for conformity to social norms. And the type of 'normativity of commitment' involved in this kind of norm-conformity is not instrumental in character.

Taking up on this idea of a social norm being a standard of evaluation for approving behaviour, the understanding of why people conform to social norms differs significantly from the notions of agency in classical economics. And this is why the understanding of the normativity or the nature of the 'ought' involved differs too. Instead of granting the agent all types of preferences as explanation for his action to find out the motivation behind norm-conformity, I argue that it is instead rather useful to analyse the normativity involved in what motivates agents to conform to social norms. I claim that we have to distinguish between two forms of normativity involved in norm-conformity: one is what I call 'instrumental normativity', which is well captured by the economic framework and the notion of preferences. The second is a kind of 'normativity of commitment', which, so the argument goes, cannot be captured by economic theory. Depending on people's motivation, they consider conformity as being normative in either one or both of the two senses at the same time. This depends on the standard of evaluation they use. If they evaluate norm-conformity in terms of appropriateness or inappropriateness, they take the shared values as a basis for their evaluation. In this case, to follow a social norm is normative for a

person, *i.e.* the person recognizes and accepts the normative claim that that social norm makes on him/her, because the norm embodies the principles and values the individual or group is committed to. And the “acceptance of a certain normative consideration can mould the way [one] thinks” (Millar, 09.05.2007, p. 1) and acts. What motivates people is indirectly the recognition of the claim the norm makes on them and directly their commitment to the principles or values embodied in the norm. Against the economists’ view, personal welfare-improvements are then not the reason for the choice of the agent; they are not what the agent aims at in the first place and thus are not what makes norm-conformity normative for them (Sen, 1977; Hausmann, 2005). In the case of norm-conformity as being an action motivated by commitment, the agent considers conformity as normative for him because of his commitment to what the social norm embodies, *i.e.* what is seen as right or wrong — and his commitment to values and principles which provides the basis for evaluation of an action to be right or wrong makes the performance of the action for him necessary, *i.e.* normative for him.

Why does this whole picture present a problem for the economist? To see this, we have to understand in detail what characterizes the motive of commitment and what it implies. The interesting issue about commitment is that it seems to matter to economists as well as to philosophers and thus perfectly represents the potential intersection of both disciplines. Whilst there is a lot of philosophical discussion going on about the nature of commitment, what it means for a person to be committed to something/someone and what commitment aims at, there are only very few economists who take the importance of commitment as a motive for agency seriously. One of these is Amartya Sen. In his article *Rational Fools: A Critique of the Behavioral Foundations of Economic Theory* published in 1977, Sen famously argues against the usefulness of the economist principle of self-interest as part of the assumption structure in economic models. He claims that not all motives for action can be reduced to ultimate self-interest. He remarks that the problem with economic theory is that “there is no choice-independent way of understanding someone’s attitude towards alternatives” (Sen, 1977, p. 323). But, he goes on to say, there is one source of action which does not involve choice in the usual sense. That source of action is the motive of commitment.

What does it mean to say that commitment is choice-independent? This means that, when motivated by commitment, an agent takes an action even though the consequences of the action will not necessarily affect his personal welfare. So, there is no choice involved in the usual sense. The person would still perform the action, even if he believes that it will yield to a negative effect on his personal welfare; he would perform it even if, at the same time, he could perform an alternative action that would make him better off. Thus, taking commitment as a possible motive for action, the identity of personal welfare and individual preferences resulting in choice no longer holds (Sen, 1977).<sup>14</sup> In the case of commitment, personal welfare and personal choice are separated from one another. As we shall see, the concept of choice is no longer based on the idea that the chosen alternative always has to be better than all other available options, *i.e.* that the choice of a rational agent would result in the optimization (or Simon's idea of satisfaction) of the person's preferences, and therefore best improves the level of personal welfare.

To introduce his idea, Sen draws on a distinction between the motive of sympathy and a motive of commitment, both are different types of other-regarding or "altruistic motivation" (Pettit, 2007). Although a 'pro-social' motive, sympathy does not require a departure from personal welfare or individual utility maximization. As shown above with the example of altruism, an action motivated by sympathy is in the same way ultimately self-regarding, which means that "the concern for others directly affects one's own welfare" (Sen, 1977, p. 326). One example for the person acting from sympathy can be taken from Kant's *Groundwork*. Kant introduced the friend of mankind, who finds pleasure in spreading joy around him and who can take delight in the satisfaction of others. But Kant sees this satisfaction as an inclination like any other desire a person can have. He concludes that such an action, even if admirable, has no moral worth (Kant, 1997). In contrast, an action which is motivated by duty has moral worth; it is characterized as an action performed independently of personal desires. Kant illustrates this point by de-

---

<sup>14</sup>This identity mainly depends on the underlying understanding of the concept of a preference or the nature of reasons. As claimed before, personal welfare and individual utility are (technically) equalled in mainstream economics. This assumption provides the main basis for conclusions about what we understand as a preference and what the nature of reasons is (Sen, 1977, 1985).

scribing a philanthropist acting kindly towards others. He does not have any inclination which would positively or negatively affect his state of affairs. He simply acts because duty demands it. According to Kant, this gives the action a genuine moral worth (*ibid.*).

Abstracting from one's personal desires and inclinations implies that the action is not driven by them. Rather, an action motivated by commitment requires a "counterpreferential choice" from the part of the agent (Sen, 1977, p. 328). That does not mean, that the agent is not allowed to have any kind of additional inclination which makes the action attractive for him. This just means that under at least one counterfactual condition the personal welfare under the act chosen would be unaffected.<sup>15</sup> Hence, the agent would unconditionally perform the action, *i.e.* he would have no choice, because his commitment makes performing the action necessary for him. And even if these inclinations or desires would not be present, he would still perform the action in question. An action motivated by commitment can even imply a choice against one's personal welfare-improvement.<sup>16</sup> Personal welfare-improvements are not the *reason* for the action of the agent; they are not what the agent aims at in the first place (Sen, 1977; Hausmann, 2005). That means, an action motivated by commitment is completely non-egoistic.

It seems obvious that there is a close relationship between commitment and moral agency. According to Kant, acting from duty requires a sense of moral commitment. Being committed to moral principles is a necessary precondition for acting from duty. Taking Kant's motive of duty as one 'subgroup' of the general idea of commitment, Sen's understanding of commitment is that an action motivated by commitment is *per definitionem* independent from the satisfaction of any kind of inclination or immediate desires the agent is driven by in a given moment (Kant, 1997). In the case of commitment,

---

<sup>15</sup>It has to be noticed that it is difficult to imagine different states of the world and potential choices for action. However, I will not discuss this difficulty further here but just present the simple idea of commitment.

<sup>16</sup>I have to admit that an underlying premise of the argument is the empirical statement that commitment as characterized in this essay does exist. However, to ground this claim in empirical evidence seems to be a difficult task. People's motivations behind their actions are often seen as black boxes. There is a lot of empirical research still to be done on filtering out whether there is such a motivating force as commitment and, if yes, to disentangle this from other motives which are active at the same time the agent performs a certain action.

as is the case with actions done from duty, the agent considers the action as the *right* thing to do. Consequently, as when acting from duty, commitment requires the agent to be capable of undertaking judgment and evaluation, being self-reflected and conscious about his values and principles upon which he bases his evaluations.<sup>17</sup>

However, although morality implies commitment, morality should not be equalled or confused with it. Commitment results in the *right* action, although this does not necessarily mean that it results in what is seen as *morally right*. A person can also be committed to an ideology, a political party, or a religious view. A person can be committed to an organisation, to a social contract, to (moral) principles or to god. *Right* in this context thus just means that the agent considers the action itself (or the outcome of the action) as worth pursuing, independent of the personal benefit he can expect from the outcome of the action. And the agent comes to know that it is *the right thing to do* because he evaluates and judges the action on the basis of the values and principles he holds while comparing possible alternative actions instead of drawing on the evaluative principle of utility-maximization. In this sense, an action motivated by commitment is, as a moral action, unconditional and stable. But it does not necessarily aim at what is the morally right thing to do.

To sum up: the idea of commitment comprises an evaluation of possible alternatives that is not necessarily carried out by basing it on moral values; instead, it can be based on any set of values the individual in question holds. However, what commitment has in common with morality, at least understood in a Kantian way, is the existence of a non-welfarist goal that is altogether removed from individual preference-satisfaction and the fulfilment of immediate inclination. And the economist does not allow for this kind of behaviour.

Are there any ways out of this explanatory dilemma? Why not just assume an unconditional preference for either a certain value or commitment *per se*? The third question will be addressed to understand more clearly why even a very open economic model has prob-

---

<sup>17</sup>The character of commitment strongly depends on the definition of concepts. Frankfurt-minded philosophers would interpret commitment in a less evaluative and a less-cognitive way. According to Frankfurt people do not always know what they are committed to. Sometimes they believe they are committed to something they actually are not for example (*e.g.* Frankfurt, 1971). I just use Sen's understanding of commitment here, which seems very much in line with a Kantian view of moral agency.

lems integrating the motive of commitment. Where is this ‘ought’ or type of normativity grounded in? Does it differ between different people/cultures and why?

### III.

One recent attempt at defining social norms while taking both types of normativity into account is by looking at people’s expectations. People’s preferences for social norms can be put in relation to normative expectations about actions that a group or the society has toward an individual or the individual has towards society (Coleman, 1987; Bicchieri, 2006). Cristina Bicchieri, one of the most prominent philosophers who work on social norms, stresses the importance of mutual expectations involved in norm-conformity. Her *Grammar of Society* published in 2006 is widely seen as “making a key contribution to our understanding of the motivations behind norm-compliance” (Küchle and Ríos, 2008, p. 118). She considers a social norm to be a kind of cluster of expectations among a sufficient number of people who believe that a social norm exists and who expect that many other people will follow it in certain situations. In this sense, people have a conditional preference for norm-conformity; conditional upon expectations about other people’s belief in whether an action is seen as being right or wrong within a community (Bicchieri, 2006).

Bicchieri’s account mainly consists of the following four (individually) necessary and (jointly) sufficient conditions for the existence of social norms: (1) People are collectively aware that a norm exists and that it applies to specific circumstances (contingency). (2) But the simple presence of the norm does not suffice to guarantee conformity. People also need to have a conditional preferences to conform to a social norm; conditional on (3) expectations of the following kind: A person believes that, based on past observations of other’s behaviour or its consequences, the majority of the people will follow the norm as well (empirical expectations) and (4) a belief that others expect a person to conform (normative expectations) in the sense that they believe that everyone ought to conform. Here, expectations are taken to be reciprocal. Normative expectations can be accompanied by sanctions but do not have to be. The people who expect others to conform and who are willing to sanction trans-

gressions also have to *prefer* people to conform. What brings in the instrumental aspect of norm-conformity is that people have to have a (conditional) preference for conformity and the required beliefs in addition to the expectations. In line with the classical belief/desire-framework of rational choice theory, people are finally motivated by their preferences (Bicchieri, 2006). What brings in the normative aspect of conformity is the condition of mutually held ‘normative expectations’ about what people ought to do, based on what people consider as right or wrong. Let us look at this condition in greater detail.

The notion of mutually held ‘normative expectations’ introduces a type of normativity that differs from the mere instrumental normativity and from having a prudential reason for conformity. It implies that people not only think that others ought to conform, but that they also believe that they themselves are expected to conform in the sense that a reasonable number of people think that one *ought* to conform to the social norm. This can suffice to induce a preference for conformity when “individuals recognize the legitimacy of others’ expectations and feel an obligation to fulfil them” (2006: 15); and, once induced, a preference for conformity guarantees that people follow the social norm. But it does not need to suffice. In some cases, Bicchieri claims, potential sanctions are needed to induce a preference for conforming to the social norm. And expecting sanction might motivate either in the sense that people fear sanctions and want to avoid them or because of a desire to please others and thus be rewarded.

What is of interest for the argument here is the nature of the ‘ought’ in Bicchieri’s account. It can be implicitly contained in the normative or the empirical belief people hold about others’ expectations. This is, people think they ought to conform because of their belief about others’ expectations. In the case of an empirical belief, people infer from my (consistent) behaviour in the past that I will do the same in the future, and that this is what I believe (Bicchieri, 2006). If it is a normative belief things look a bit different. Although normative beliefs do not generally state an obligation, they do so in Bicchieri’s account (*ibid.*). In this case, I believe that a sufficiently large number of people think that I have an obligation to conform to a social norm in the appropriate circumstances (*ibid.*). This is why a person will consider other people’s expectations as legitimate.



It is not a kind of instrumental normativity, because the motivation for conformity understood in this sense can but does not have to incur material or immaterial sanctions to guarantee conformity. Also, conforming to the norm does not simply serve as a means to satisfy my ends. Instead, people consider conformity to the social norm as their and other people's obligation or duty. But why does the person accept this obligation and thus considers other people's expectations as legitimate? Bicchieri claims that people think that they (and others) are obliged to conform to the norm because they believe it is the right thing to do. In these cases, social norms are seen as good or reasonable.<sup>18</sup> But how do people justify this obligation or, to put it differently, what is this obligation grounded upon? Bicchieri does not give a final answer. But she indicates that we have a preference to share a cake equally for example because it might be our duty to be fair. And we hold the belief that others expect us that we ought to share because it is the fair thing to do. But on what grounds do people believe that the norm of equal sharing is the right thing to do?

We could get around this problem by arguing that social norms are often not seen to be the right thing to do. But then the accepted obligation involved seems even more difficult to understand. To make her approach understandable, Bicchieri gives the following example of a Muslim woman wearing a veil: although it is a widely spread social norm, many people (in the Muslim community) might think that wearing a veil is an unpleasant thing to do and not many people would be prepared to sanction a transgression. This means, people would not consider it an obligation to conform to the norm. Each woman therefore does not necessarily believe that she ought to follow the social norm. But she might believe that she is expected to wear the veil in the sense that she believes that many people think that she ought to wear a veil and that they also prefer to wear a veil because it is an obligation or her religious duty to wear it. Because of social pressure she will finally end up wearing it; expected sanctions are necessary to guarantee conformity. Her reason to conform

---

<sup>18</sup>Of course, as Bicchieri rightly stresses, not all social norms are thought of being good and people do not think that others ought to follow them. Yet, conformity to them is widely observed. I will not consider this case here, but I am aware that this is an important issue which has to be stressed, when norm-conformity should be explained fully.

is thus a kind of ‘instrumental reason’ and the type of normativity involved has an instrumental nature. But let us look at this case in greater detail: what this case implies is that a religious community of extensive size would follow a norm which dramatically constrains the ‘quality of life’ of at least half of the members of the community, whereas ‘quality of life’ labels a lifestyle taken from the idea of freedom of secular western societies. And it assumes that people accept this constraint on their choices only because of other people’s expectations, which, even worse, is based on false beliefs because, as Bicchieri describes, these expectations do not really exist if we would ask people for people’s actual beliefs. What this example suggests, then, is that people live against what they really or truly value and consider as the right thing to do, simply because of a falsely believed existence of expectations.

This seems unreasonable to me. I guess that a vast majority of Muslim women would wear the veil even if they would not fear any kind of sanctions. And this is because they do it out of a religious conviction. They have committed themselves to the values and principles of their religious community. And they believe that conformity to the social norm is one way of expressing this conviction. This commitment is why they think they ought to follow the norm and why they also expect others to conform. Such a view immediately loses its validity when people have different values. Western women would have difficulty in believing in this kind of ‘uncomfortable’ commitment. Because of their commitment to other kind of values, their motivation and the normativity involved would be different immediately, *i.e.* rather instrumental in character. So, the conformity to the same norm can be motivated in at least two different ways, namely either fear of sanctions or conviction and commitment to values and principles that are embodied in the norm. Consequently, why people think they ought to follow a social norm also differs. As the one is instrumental in character and guarantees conformity only by expected sanctions, the conformity of a person who is motivated in this way is *contingent*. As the other one is ‘committing’ in character and guarantees conformity also without any kind of expected sanctions (the person conforms voluntarily and unconditionally), the conformity of a person who is motivated in this way is *necessary*.

I think this relates to an idea Hume raised (Hume, 2003). Sometimes norms and values or virtues appear to be related. Bicchieri

claims that once norms are invoked in a society, people start to follow them because they attribute some value to what the norm stands for or they have a deep conviction of the norm. Consequently, the norm is what shapes the values of the people within a community (Bicchieri, 2006). Accordingly, when we attribute an intrinsic value to the norm, this gives us a reason to recognize other people's expectations as legitimate and we feel an obligation to fulfil it. These expectations cause us to have a preference and having a preference causes us to conform to the norm we value or are deeply convinced of. So their 'value depends on widespread conformity' (Bicchieri, 2006). I think this seems somewhat contradictory. We consider an action worth pursuing because its properties embody certain values and principles we already have or hold and about which we are convinced. We are committed to these values and, by following a norm, we want to implement or realize them. Consequently, it is often our values and principles that shape social norms, not the other way round.<sup>19</sup> Because we are committed to the principles and values we hold, we take it as our obligation to conform to the norm. This obligation is independent, first, from our preferences and thus from what is in our self-interest and, second, from other people's expectations.

So, as I claimed before, I think that individuals draw on a basis of values and principles that people commonly hold. And because they commit themselves to these values and principles, on the one hand, they accept other people's evaluations and expectations about what kind of behaviour is considered as appropriate and find them legitimate. On the other hand, their commitment also justifies their own expectations towards other people's behaviour and their own norm-compliance. This, so the economist will argue, seems vague and does not explain properly the normativity involved. We should rather grant people to have preferences for fairness or for other social concerns which would replace this talk about values. And, if we

---

<sup>19</sup>Empirically it is probably true that in many societies certain social norms exist before new individuals build up values that support these norms. But what was there first and how (informal) institutions evolve are question which cannot be addressed in this paper. However, I think that, if what the social norms represent or imply would be contradictory to what we consider for ourselves as acceptable and would not be more or less inline with our values, conceptions, and ideas about the world, the way we live together in a society, and how we think we should treat each other, then the majority of the people would not conform to social norms.

cannot get around the idea of commitment, we could simply assume them to have a preference for commitment.

But what would a preference for commitment look like? Having characterized what the nature of commitment is, it seems plausible that the idea of a preference and what it implies — both technically and factually — seems to be misleading in this context. Commitment is inherently contradictory to a self-interested action and does not lead to personal-welfare improvement in the first place (Sen, 1985, 2005). One could object that Bicchieri allows for counter-preferential choices: in a situation where (material) rewards are present, I might prefer more to less but when people expect me to follow a norm of generosity ‘I might prefer to behave generously’. Is this what we intuitively understand as counter-preferential choice? It rather sounds like having a dominant preference for conforming to the norm of generosity; we can change our preference or one preference dominates the other, but then we do not take a choice which contrasts our preferences. Thus, the action is still based on and not against our preferences.

This relates to the point that the dominant role of expectations and the way in which we form our preference for social norms seems quite odd. Bicchieri unifies preference-formation and the existence of expectations and does not focus on whether other’s expectations make any sense for us. In the long run, it would be impossible for an agent to keep his promises in a world where nobody kept theirs. He would have to change his preference in accordance with the change of his beliefs about other’s expectations, independent of whether he cares about taking the specific action ‘keeping a promise’, *i.e.* whether he judges the action of keeping promises as good or the right thing to do. It does not matter what kind of behaviour we believe to be expected, it could be keeping a promise or it could be jumping as high as we can when we see a green bird. But — according to the theory — whether we consider the action to follow a norm worth pursuing or the right thing to do (which is obviously not the same as having a preference) does not matter.<sup>20</sup> This conflicts with the idea of commitment. But why would people prefer what other people expect them to do, no matter how ‘stupid’ it seems to

---

<sup>20</sup>Note that on this account it is very likely that norm-guided behaviour will change with circumstances, because expectations may be different in distinct contexts, which changes the preference of the agent and thus his behaviour.

be? Why would all Muslim women accept to be strongly limited in their personal freedom, if they completely disagreed with what the norm says and even formed a preference to this extent?<sup>21</sup> Even if this point appeals only to our intuition and thus has yet to be proven, I doubt that any expectations we believe people to hold have this power to form people's preferences accordingly.

This leads us to the last point, namely that awareness, self-reflexion and conscious deliberation are not necessary conditions for following a norm in Bicchieri's account. She claims that we often unconsciously follow social norms, while she defines a mental state to be conscious when it is accompanied by roughly simultaneous, higher-order thought about that very mental state. This understanding of awareness is reminiscent of conscious deliberation modelled in rational choice theory. In a cost-benefit-analysis, we consciously weigh our options available. In this context, deliberation is based on the beliefs and desires of which we are aware in the abovementioned sense. I believe that Bicchieri understates the importance of this kind of awareness and claims that we often make use of heuristics to 'choose' our actions, especially when it comes to norm-conformity. And she might be right in many cases. However, in the context of norm-conformity motivated by commitment, awareness also requires making judgements and it implies self-consciousness achieved by a process of deliberating about how we act on the whole. We draw on our experience of how we acted in the past and why we acted as we did to make sure that our action was worth pursuing and to adapt/improve our action in the future, having in mind certain fundamental values and principles to which we commit ourselves. As said before, because of this commitment, people use their values and principles for assessment. These values and principles for assessment are different from the notion of value implied by Bicchieri's claim that some social norms may become part of our system of values, and that we may feel a strong obligation to obey them (Bicchieri, 2006). In-

---

<sup>21</sup>Of course, the economist would argue that Bicchieri's account is just a rational reconstruction and does not aim at showing how people's *real* preferences are (Bicchieri, 2006). But besides the fact, that a rational reconstruction of an account for norm-conformity seems problematic as people, depending on their motives, conform to social norms in a more or less stable manner, it seems problematic to talk about what shapes people's preferences in the first place, when we just assume that people have preferences. In this case the whole talk about expectations would lose importance.

stead, as mentioned before, the concept of commitment implies that some norm-governed actions embody people's values and principles. And because the norm embodies the values and principles we are committed to, it has a normative claim on us which is a general claim independent of the situation. So, once we judge an action as worth pursuing on the basis of our values and principles, we have the duty to pursue it; in case of social norms 'we obey ourselves' to conform. The awareness of the existence of these properties, of the fact that they provide us with a reason to act, and of the values and principles we hold is therefore crucial when we want to account for the motive of commitment.

Economists might argue that, if we do not have a preference for the norm, these values constrain us in following our desires, so why should we commit ourselves to them? Let me give a constructivist answer to that question: these assessment criteria are constructed by the agents themselves through their own will (Korsgaard, 1997). During the process of deliberation, where the agent engages in critical thinking, he enlarges his perspective and option-set and gets a deeper understanding of the situation and of himself (Wallace, 2003). He becomes conscious of values and principles on which he wants to base his actions. This process does change his intentions; it is a kind of self-determination. Instead of passively reacting to his desires, he deliberates and assesses his own reasons for his actions on the basis of these values and principles. And because they reflect his overall view of what is worth to pursue in life, he wills them to have a claim on him.

To sum up, Bicchieri's definition of social norms has the advantage of being operational. It provides us with a pragmatic framework. It has empirical, testable content which allows us to predict norm-governed behaviour, especially in experimental settings. At the same time it is highly flexible and sophisticated enough to include aspects like emotions and internalization of norms. However, while it is able to explain nearly all types of motives people can have for norm-conformity, it also shows that an action motivated by commitment cannot be explained. The only way out would be to induce a preference for commitment but this is not what we are looking for, as personal welfare-improvement and self-interest are not part of the motive of commitment. Hence, it is problematic to account for norm-conformity in a satisfactory way with an approach where choice is

based on the motive of (immediate) satisfaction of individual desires or preferences and thereby aiming at personal welfare-improvements.

#### **IV.**

It has to be analysed further whether the theory of human behaviour used in economics to explain norm-conformity fulfils its explanatory and predictive role when accounting for norm-conformity. It has to be asked how useful economic explanations of norm-conformity are. To enhance their usefulness and adequacy on the one hand, and be more realistic in what they predict on the other, we have to understand what really motivates people. And finding out the motivation, the normativity involved in the conformity of the people, plays a crucial role. Thus, analysing the normativity involved would help us to improve our explanations in economics and explain cases of cooperation which are just ad hoc explained by standard as-if-models based on the rational choice approach. This difference in motivation and thus in the normativity involved has two important consequences. First, it has important implications on a methodological level when giving an explanation of norm-conformity. The explanation based on an account where normativity is assumed to be purely instrumental seems to be misplaced, as it ignores the possibility for a normativity of a different type. Second, it has important implications on a conceptual level. For, what we define as a social norm in contrast to other rules depends on why people think they ought to follow the social norm, *i.e.* what motivates their conformity. And if some motives can be covered but others are left out we can have legitimate doubts about the usefulness of the explanation provided.

My guess is that there are two ways out of this ‘explanatory dilemma’. Either, we fully dismiss economic models when it comes to explaining norm-conformity. Or, we conduct further theoretical and empirical research in this specific area to extend the economic model, improve our utility functions and provide better cognitive and behavioural assumptions of the agents.

#### **IV Conclusion**

The aim of the paper was to show the difficulty economists have in accommodating the motive of commitment in their explanatory frameworks for norm-conformity. This is because in the case of norm-

conformity, the normativity involved can be instrumental and/or committing. Of course, people's behaviour is complex and a theoretical framework will probably never cover all of the potential motivations for following social norms. However, I conclude that if we aim at reasonable explanations of norm-conformity, all possible motives should be considered. The concept of preferences is often misleading when it comes to explaining norm-conformity, as it dismisses the normative dimension of agency, *i.e.* the ability of the agent to undertake judgements and make choices independently of one's personal desires or inclination, *i.e.* independent of what is in the person's self-interest. Instead of granting the agent all types of preferences, to find out the motivation behind norm-conformity, it is rather useful to analyse the normativity involved in what motivates agents to conform to social norms. The type of normativity or the 'ought' involved in social norms differs remarkably, depending on the motivation people have. I argued that we have to distinguish between two forms of normativity involved in norm-conformity: one is what I called 'instrumental normativity', which is well captured by the economic framework and the notion of preferences. The second is a kind of 'normativity of commitment', which, so the argument was, cannot be captured by economic theory. Depending on people's motivation, they consider conformity as normative either in one of the two senses or in both at the same time. The main difficulty of the economic theory to account for these aspects lies in its general understanding of what is meant when we say that somebody acts for a reason and what we consider as being rational within the framework. It does not account either for the kind of reason we are provided with when acting from commitment, nor for what we consider as a rational in this context. Thus, if we accept commitment as a motivational source, then we must challenge our notion of reasons, *i.e.* the nature of the desires and beliefs involved in reasons for conformity and our understanding of what it means to act rationally in this context.

### References

- Akerlof, G. A., 1982. Labor contracts as partial gift exchange. *Quarterly Journal of Economics* 97 (4), 543–569.
- Anderson, E., 2000. Beyond homo economicus: New developments in theories of social norms. *Philosophy and Public Affairs* 29 (2),



170–200.

Anscombe, G. E. M., 1963. *Intention*, 2nd Edition. Basil Blackwell, Oxford.

Basu, K., Weibull, J. W., 2003. Punctuality: A cultural trait as equilibrium. In: Arnott, R. J., Greenwald, B. C. N., Kanbur, S. M. R., Nalebuff, B. J. (Eds.), *Economics for an imperfect world: Essays in honor of Joseph E. Stiglitz*. MIT Press, Cambridge, MA, pp. 163–182.

Bicchieri, C., 2006. *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press, Cambridge.

Coleman, J. S., 1987. Norms as social capital. In: Radnitzky, G., Bernholz, P. (Eds.), *Economic imperialism: The economic approach applied outside the field of economics*. A Professors World Peace Academy Book. Paragon House, New York, NY.

Coleman, J. S., 1990. *Foundations of social theory*, 1st Edition. Belknap Press of Harvard University Press, Cambridge, MA [a.o.].

Conlin, M., Lynn, M., O'Donoghue, T., 2003. The norm of restaurant tipping. *Journal of Economic Behavior & Organization* 52 (3), 297–321.

Davidson, D., 1963. Actions, reasons, causes. *Journal of Philosophy* 60 (23), 685–700.

Davis, J. B., 2004. Identity and commitment: Sen's conception of the individual. Tech. Rep. TI 2004-055/2, Tinbergen Institute, Amsterdam.

Dawes, R., Thaler, R., 1988. Anomalies: Cooperation. *Journal of Economic Perspectives* 2 (3), 187–197.

Elster, J., 1989a. Social norms and economic theory. *Journal of Economic Perspectives* 3 (4), 99–117.

Elster, J., 1989b. *Solomonic judgements: Studies in the limitations of rationality*. Cambridge University Press, Cambridge [a.o.].

- Elster, J., 1994. The nature and scope of rational-choice explanation. In: Martin, M., McIntyre, L. C. (Eds.), *Readings in the philosophy of social science*, 2nd Edition. A Bradford book. MIT Press, Cambridge, MA, pp. 311–322.
- Fehr, E., Fischbacher, U., 2003. The nature of human altruism. *Nature* 425, 785–791.
- Fehr, E., Fischbacher, U., 2004. Social norms and human cooperation. *TRENDS in Cognitive Sciences* 8 (4), 185–190.
- Fehr, E., Kirchler, E., Weichbold, A., Gächter, S., 1998. When social norms overpower competition: gift exchange in experimental labor markets. *Journal of Labor Economics* 16 (2), 324–351.
- Frankfurt, H. G., 1971. Freedom of the will and the concept of a person. *The Journal of Philosophy* 68 (1), 5–20.
- Hausman, D. M., McPherson, M. S., 2006. *Economic analysis, moral philosophy, and public policy*, 2nd Edition. Cambridge University Press, Cambridge.
- Hausmann, D. M., 2005. Sympathy, commitment, and preference. *Economics and Philosophy* 21, 33–50.
- Heap Hargreaves, S., Hollis, M., Sugden, R., Weale, A., 1992. *Theory of choice: A critical guide*. Blackwell, Oxford.
- Hechter, M., Opp, K.-D. (Eds.), 2001. *Social norms*. Russell Sage Foundation, New York, NY.
- Henrich, J., 2000. Does culture matter in economic behaviour? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. *The American Economic Review* 90 (4), 973–979.
- Hume, D., 2003. *A treatise of human nature*. Dover Publications, Mineola N.Y.
- Kahneman, D., Knetsch, J. L., Thaler, R., 1986. Fairness as a constraint on profit seeking: entitlements in the market. *The American Economic Review* 76 (4), 728–741.

- Kant, I., 1997. Groundwork of the metaphysics of morals. Cambridge texts in the history of philosophy. Cambridge University Press, Cambridge.
- Kerr, N. L., Garst, J., Lewandowski, D. A., Harris, S. E., 1997. That still, small voice: commitment to cooperate as an internalized versus a social norm. *Personality and Social Psychology Bulletin* 23, 1300–1311.
- Kirchgässner, G., 2006. Das ökonomische Verhaltensmodell: Der Homo Oeconomicus. In: Nell, V. v., Kufeld, K. (Eds.), *Homo oeconomicus: Ein neues Leitbild in der globalisierten Welt?* Vol. 1 of *Forum Philosophie & Wirtschaft*. LIT Verlag, Berlin, Hamburg, Münster, pp. 81–107.
- Korsgaard, C. M., 1997. The normativity of instrumental reason. In: Cullity, G., Gaut, B. (Eds.), *Ethics and practical reason*. Clarendon Press, Oxford, pp. 215–254.
- Krupka, E. L., Weber, R. A., 2008. Identifying social norms using coordination games: Why does dictator game sharing vary? Tech. Rep. 3860, Institute for the Study of Labor (IZA), Bonn, available online from <http://ftp.iza.org/dp3860.pdf>, last accessed: 24.02.2009.
- Küchle, G., Ríos, D., 2008. Book review of Cristina Bicchieri's *Grammar of society*. *Economics and Philosophy* 24 (1), 117–123.
- Kurzban, R., Houser, D., 2005. Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences of the United States of America* 102 (2), 1803–1807.
- Lindbeck, A., Nyberg, S., Weibull, J. W., 1999. Social norms and economic incentives in the welfare state. *The Quarterly Journal of Economics* 114 (1), 1–35.
- Mantzavinos, C. (Ed.), 2001. *Individuals, institutions, and markets. Political economy of institutions and decisions*. Cambridge University Press, Cambridge.
- Millar, A., 09.05.2007. The explanatory role of normative considerations. Available online from <http://www.>

[essex.ac.uk/philosophy/tpn/documents/Normativity/](http://essex.ac.uk/philosophy/tpn/documents/Normativity/MillarExplanatoryRole.pdf)  
[MillarExplanatoryRole.pdf](http://essex.ac.uk/philosophy/tpn/documents/Normativity/MillarExplanatoryRole.pdf).

- Okruch, S., 1999. Innovation und Diffusion von Normen: Grundlagen und Elemente einer evolutorischen Theorie des Institutionenwandels. Vol. 491 of *Volkswirtschaftliche Schriften*. Duncker & Humblot, Berlin.
- Opp, K.-D., 1999. Contending conceptions of the theory of rational action. *Journal of Theoretical Politics* 11 (2), 171–202.
- Pauer-Studer, H., 2007. Instrumental rationality versus practical reason: desires, ends, and commitment. In: Peter and Schmid (2007), pp. 73–104.
- Peter, F., Schmid, H. B. (Eds.), 2007. *Rationality and commitment*. Oxford University Press, Oxford.
- Pettit, P., 1990. The reality of rule-following. *Mind* 99 (393), 1–21.
- Pettit, P., 2007. Construing Sen on commitment. In: Peter and Schmid (2007), pp. 28–49.
- Raz, J., 2005. The myth of instrumental rationality. *Journal for Ethics and Social Philosophy* 1 (1), 1–28.
- Rosenberg, A., 1995. *Philosophy of social science*, 2nd Edition. Dimensions of philosophy series. Westview Press, Boulder.
- Samuelson, P. A., 1938. A note on the pure theory of consumer's behavior. *Economica* 5 (17), 61–71.
- Schmid, H. B., 2007. Beyond self-goal choice: Amartya Sen's analysis of the structure of commitment and the role of shared desires. In: Peter and Schmid (2007), pp. 211–226.
- Searle, J. R., 2001. *Rationality in action*. Vol. 2000 of *A Bradford book*. MIT Press, Cambridge, MA.
- Sen, A. K., 1977. Rational fools: A critique of the behavioral foundations of economic theory. *Philosophy and Public Affairs* 6 (4), 314–344.

- Sen, A. K., 1982. Choice, welfare and measurement. MIT Press, Cambridge, MA.
- Sen, A. K., 1985. Goals, commitment, and identity. *Journal of Law, Economics, and Organization* 1 (2), 341–355.
- Sen, A. K., 2005. Why exactly is commitment important for rationality? *Economics and Philosophy* 21, 5–13.
- Stout, L. A., 2001. Other-regarding preferences and social norms. Tech. Rep. Georgetown Law and Economics Research Paper No. 265902; and Georgetown Public Law Research Paper No. 265902, Georgetown University Law Center, Washington, DC, available online from [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=265902](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=265902), last accessed: 22.02.2009.
- Sugden, R., 2000. Team preferences. *Economics and Philosophy* 16 (175-204).
- Turnbull, C. M., 1972. The mountain people. Simon & Schuster, New York.
- Wallace, R. J., 2003. Practical reason. In: Zalta, E. N. (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2003 Edition).
- Wallace, R. J., Smith, M., Pettit, P. (Eds.), 2004. Reason and value: Themes from the moral philosophy of Joseph Raz, 1st Edition. Oxford University Press, Oxford.
- Woodward, J., 2009. Why do people cooperate as much as they do. In: Mantzavinos, C. (Ed.), *Philosophy of the social sciences: Philosophical theory and scientific practice*. Cambridge University Press, Cambridge, pp. 295–357.